



OIPE

ENTERED

RAW SEQUENCE LISTING

DATE: 07/08/2002

PATENT APPLICATION: US/10/033,325

TIME: 11:12:54

Input Set : A:\GC502-2-C1-SEQLIST.txt

Output Set: N:\CRF3\07082002\J033325.raw

```

3 <110> APPLICANT: Poulouse, Ayrookaran J.
4   Schellenberger, Volker
5   Kellis, Jr., James T.
6   Paech, Christian
7   Nadherny, Joanne
8   Naki, Donald P.
9   Collier, Katherine D.
10  Caldwell, Robert M.
11  Baeck, Andre C.
13 <120> TITLE OF INVENTION: Multiply-Substituted Protease Variants
15 <130> FILE REFERENCE: GC502-2-C1
17 <140> CURRENT APPLICATION NUMBER: US 10/033,325
18 <141> CURRENT FILING DATE: 2001-11-02
20 <150> PRIOR APPLICATION NUMBER: US 09/178,155
21 <151> PRIOR FILING DATE: 1998-10-23
23 <150> PRIOR APPLICATION NUMBER: US 08/956,323
24 <151> PRIOR FILING DATE: 1997-10-23
26 <150> PRIOR APPLICATION NUMBER: US 08/956,564
27 <151> PRIOR FILING DATE: 1997-10-23
29 <150> PRIOR APPLICATION NUMBER: US 08/956,324
30 <151> PRIOR FILING DATE: 1997-10-23
32 <160> NUMBER OF SEQ ID NOS: 6
34 <170> SOFTWARE: FastSEQ for Windows Version 3.0
36 <210> SEQ ID NO: 1
37 <211> LENGTH: 1497
38 <212> TYPE: DNA
39 <213> ORGANISM: Bacillus amyloliquefaciens
41 <220> FEATURE:
42 <221> NAME/KEY: CDS
43 <222> LOCATION: (96)...(1245)
45 <400> SEQUENCE: 1
46 ggtctactaa aatattattc catactatac aattaatata cagaataatc tgtctattgg      60
47 ttattctgca aatgaaaaaa aggagaggat aaaga gtg aga ggc aaa aaa gta      113
48                               Met Arg Gly Lys Lys Val
49                               1           5
51 tgg atc agt ttg ctg ttt gct tta gcg tta atc ttt acg atg gcg ttc      161
52 Trp Ile Ser Leu Leu Phe Ala Leu Ala Leu Ile Phe Thr Met Ala Phe
53          10           15           20
55 ggc agc aca tcc tct gcc cag gcg gca ggg aaa tca aac ggg gaa aag      209
56 Gly Ser Thr Ser Ser Ala Gln Ala Ala Gly Lys Ser Asn Gly Glu Lys
57          25           30           35
59 aaa tat att gtc ggg ttt aaa cag aca atg agc acg atg agc gcc gct      257
60 Lys Tyr Ile Val Gly Phe Lys Gln Thr Met Ser Thr Met Ser Ala Ala

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/033,325

DATE: 07/08/2002

TIME: 11:12:54

Input Set : A:\GC502-2-C1-SEQLIST.txt

Output Set: N:\CRF3\07082002\J033325.raw

61	40	45	50	
63	aag aag aaa gat gtc att tct gaa aaa ggc ggg aaa gtg caa aag caa	305		
64	Lys Lys Lys Asp Val Ile Ser Glu Lys Gly Gly Lys Val Gln Lys Gln			
65	55 60 65 70			
67	ttc aaa tat gta gac gca gct tca gtc aca tta aac gaa aaa gct gta	353		
68	Phe Lys Tyr Val Asp Ala Ala Ser Val Thr Leu Asn Glu Lys Ala Val			
69	75 80 85			
71	aaa gaa ttg aaa aaa gac ccg agc gtc gct tac gtt gaa gaa gat cac	401		
72	Lys Glu Leu Lys Lys Asp Pro Ser Val Ala Tyr Val Glu Glu Asp His			
73	90 95 100			
75	gta gca cat gcg tac gcg cag tcc gtg cct tac ggc gta tca caa att	449		
76	Val Ala His Ala Tyr Ala Gln Ser Val Pro Tyr Gly Val Ser Gln Ile			
77	105 110 115			
79	aaa gcc cct gct ctg cac tct caa ggc tac act gga tca aat gtt aaa	497		
80	Lys Ala Pro Ala Leu His Ser Gln Gly Tyr Thr Gly Ser Asn Val Lys			
81	120 125 130			
83	gta gcg gtt atc gac agc ggt atc gat tct tct cat cct gat tta aag	545		
84	Val Ala Val Ile Asp Ser Gly Ile Asp Ser Ser His Pro Asp Leu Lys			
85	135 140 145 150			
87	gta gca agc gga gcc agc atg gtt cct tct gaa aca aat cct ttc caa	593		
88	Val Ala Ser Gly Ala Ser Met Val Pro Ser Glu Thr Asn Pro Phe Gln			
89	155 160 165			
91	gac aac aac tct cac gga act cac gtt gcc ggc aca gtt gcg gct ctt	641		
92	Asp Asn Asn Ser His Gly Thr His Val Ala Gly Thr Val Ala Ala Leu			
93	170 175 180			
95	aat aac tca atc ggt gta tta ggc gtt gcg cca agc gca tca ctt tac	689		
96	Asn Asn Ser Ile Gly Val Leu Gly Val Ala Pro Ser Ala Ser Leu Tyr			
97	185 190 195			
99	gct gta aaa gtt ctc ggt gct gac ggt tcc ggc caa tac agc tgg atc	737		
100	Ala Val Lys Val Leu Gly Ala Asp Gly Ser Gly Gln Tyr Ser Trp Ile			
101	200 205 210			
103	att aac gga atc gag tgg gcg atc gca aac aat atg gac gtt att aac	785		
104	Ile Asn Gly Ile Glu Trp Ala Ile Ala Asn Asn Met Asp Val Ile Asn			
105	215 220 225 230			
107	atg agc ctc ggc gga cct tct ggt tct gct gct tta aaa gcg gca gtt	833		
108	Met Ser Leu Gly Gly Pro Ser Gly Ser Ala Ala Leu Lys Ala Ala Val			
109	235 240 245			
111	gat aaa gcc gtt gca tcc ggc gtc gta gtc gtt gcg gca gcc ggt aac	881		
112	Asp Lys Ala Val Ala Ser Gly Val Val Val Val Ala Ala Ala Gly Asn			
113	250 255 260			
115	gaa ggc act tcc ggc agc tca agc aca gtg ggc tac cct ggt aaa tac	929		
116	Glu Gly Thr Ser Gly Ser Ser Ser Thr Val Gly Tyr Pro Gly Lys Tyr			
117	265 270 275			
119	cct tct gtc att gca gta ggc gct gtt gac agc agc aac caa aga gca	977		
120	Pro Ser Val Ile Ala Val Gly Ala Val Asp Ser Ser Asn Gln Arg Ala			
121	280 285 290			
123	tct ttc tca agc gta gga cct gag ctt gat gtc atg gca cct ggc gta	1025		
124	Ser Phe Ser Ser Val Gly Pro Glu Leu Asp Val Met Ala Pro Gly Val			
125	295 300 305 310			

RAW SEQUENCE LISTING

DATE: 07/08/2002

PATENT APPLICATION: US/10/033,325

TIME: 11:12:54

Input Set : A:\GC502-2-C1-SEQLIST.txt

Output Set: N:\CRF3\07082002\J033325.raw

```

127 tct atc caa agc acg ctt cct gga aac aaa tac ggg gcg tac aac ggt      1073
128 Ser Ile Gln Ser Thr Leu Pro Gly Asn Lys Tyr Gly Ala Tyr Asn Gly
129          315          320          325
131 acg tca atg gca tct ccg cac gtt gcc gga gcg gct gct ttg att ctt      1121
132 Thr Ser Met Ala Ser Pro His Val Ala Gly Ala Ala Ala Leu Ile Leu
133          330          335          340
135 tct aag cac ccg aac tgg aca aac act caa gtc cgc agc agt tta gaa      1169
136 Ser Lys His Pro Asn Trp Thr Asn Thr Gln Val Arg Ser Ser Leu Glu
137          345          350          355
139 aac acc act aca aaa ctt ggt gat tct ttg tac tat gga aaa ggg ctg      1217
140 Asn Thr Thr Thr Lys Leu Gly Asp Ser Leu Tyr Tyr Gly Lys Gly Leu
141          360          365          370
143 atc aac gta caa gcg gca gct cag taa a acataaaaaa ccggccttgg      1265
144 Ile Asn Val Gln Ala Ala Ala Gln *
145 375          380
147 cccgcgcggt tttttattat tttttcttct ccgcatgttc aatccgctcc ataatcgacg      1325
148 gatggctccc tctgaaaatt ttaacgagaa acggcggtt gaccggctc agtcccgtaa      1385
149 cggccaactc ctgaaacgtc tcaatcgccg cttcccggtt tccggtcagc tcaatgccat      1445
150 aacggtcggc ggcgttttcc tgataccggg agacggcatt cgtaatcgga tc      1497
152 <210> SEQ ID NO: 2
153 <211> LENGTH: 382
154 <212> TYPE: PRT
155 <213> ORGANISM: Bacillus amyloliquefaciens
157 <400> SEQUENCE: 2
158 Met Arg Gly Lys Lys Val Trp Ile Ser Leu Leu Phe Ala Leu Ala Leu
159 1          5          10          15
160 Ile Phe Thr Met Ala Phe Gly Ser Thr Ser Ser Ala Gln Ala Ala Gly
161          20          25          30
162 Lys Ser Asn Gly Glu Lys Lys Tyr Ile Val Gly Phe Lys Gln Thr Met
163          35          40          45
164 Ser Thr Met Ser Ala Ala Lys Lys Lys Asp Val Ile Ser Glu Lys Gly
165          50          55          60
166 Gly Lys Val Gln Lys Gln Phe Lys Tyr Val Asp Ala Ala Ser Val Thr
167          65          70          75          80
168 Leu Asn Glu Lys Ala Val Lys Glu Leu Lys Lys Asp Pro Ser Val Ala
169          85          90          95
170 Tyr Val Glu Glu Asp His Val Ala His Ala Tyr Ala Gln Ser Val Pro
171          100          105          110
172 Tyr Gly Val Ser Gln Ile Lys Ala Pro Ala Leu His Ser Gln Gly Tyr
173          115          120          125
174 Thr Gly Ser Asn Val Lys Val Ala Val Ile Asp Ser Gly Ile Asp Ser
175          130          135          140
176 Ser His Pro Asp Leu Lys Val Ala Ser Gly Ala Ser Met Val Pro Ser
177          145          150          155          160
178 Glu Thr Asn Pro Phe Gln Asp Asn Asn Ser His Gly Thr His Val Ala
179          165          170          175
180 Gly Thr Val Ala Ala Leu Asn Asn Ser Ile Gly Val Leu Gly Val Ala
181          180          185          190
182 Pro Ser Ala Ser Leu Tyr Ala Val Lys Val Leu Gly Ala Asp Gly Ser

```

RAW SEQUENCE LISTING

DATE: 07/08/2002

PATENT APPLICATION: US/10/033,325

TIME: 11:12:54

Input Set : A:\GC502-2-C1-SEQLIST.txt

Output Set: N:\CRF3\07082002\J033325.raw

```

183          195          200          205
184 Gly Gln Tyr Ser Trp Ile Ile Asn Gly Ile Glu Trp Ala Ile Ala Asn
185          210          215          220
186 Asn Met Asp Val Ile Asn Met Ser Leu Gly Gly Pro Ser Gly Ser Ala
187 225          230          235          240
188 Ala Leu Lys Ala Ala Val Asp Lys Ala Val Ala Ser Gly Val Val Val
189          245          250          255
190 Val Ala Ala Ala Gly Asn Glu Gly Thr Ser Gly Ser Ser Ser Thr Val
191          260          265          270
192 Gly Tyr Pro Gly Lys Tyr Pro Ser Val Ile Ala Val Gly Ala Val Asp
193          275          280          285
194 Ser Ser Asn Gln Arg Ala Ser Phe Ser Ser Val Gly Pro Glu Leu Asp
195          290          295          300
196 Val Met Ala Pro Gly Val Ser Ile Gln Ser Thr Leu Pro Gly Asn Lys
197 305          310          315          320
198 Tyr Gly Ala Tyr Asn Gly Thr Ser Met Ala Ser Pro His Val Ala Gly
199          325          330          335
200 Ala Ala Ala Leu Ile Leu Ser Lys His Pro Asn Trp Thr Asn Thr Gln
201          340          345          350
202 Val Arg Ser Ser Leu Glu Asn Thr Thr Thr Lys Leu Gly Asp Ser Leu
203          355          360          365
204 Tyr Tyr Gly Lys Gly Leu Ile Asn Val Gln Ala Ala Ala Gln
205          370          375          380
207 <210> SEQ ID NO: 3
208 <211> LENGTH: 275
209 <212> TYPE: PRT
210 <213> ORGANISM: Bacillus amyloliquefaciens
212 <400> SEQUENCE: 3
213 Ala Gln Ser Val Pro Tyr Gly Val Ser Gln Ile Lys Ala Pro Ala Leu
214 1          5          10          15
215 His Ser Gln Gly Tyr Thr Gly Ser Asn Val Lys Val Ala Val Ile Asp
216          20          25          30
217 Ser Gly Ile Asp Ser Ser His Pro Asp Leu Lys Val Ala Gly Gly Ala
218          35          40          45
219 Ser Met Val Pro Ser Glu Thr Asn Pro Phe Gln Asp Asn Asn Ser His
220          50          55          60
221 Gly Thr His Val Ala Gly Thr Val Ala Ala Leu Asn Asn Ser Ile Gly
222 65          70          75          80
223 Val Leu Gly Val Ala Pro Ser Ala Ser Leu Tyr Ala Val Lys Val Leu
224          85          90          95
225 Gly Ala Asp Gly Ser Gly Gln Tyr Ser Trp Ile Ile Asn Gly Ile Glu
226          100          105          110
227 Trp Ala Ile Ala Asn Asn Met Asp Val Ile Asn Met Ser Leu Gly Gly
228          115          120          125
229 Pro Ser Gly Ser Ala Ala Leu Lys Ala Ala Val Asp Lys Ala Val Ala
230          130          135          140
231 Ser Gly Val Val Val Val Ala Ala Ala Gly Asn Glu Gly Thr Ser Gly
232 145          150          155          160
233 Ser Ser Ser Thr Val Gly Tyr Pro Gly Lys Tyr Pro Ser Val Ile Ala

```

RAW SEQUENCE LISTING

DATE: 07/08/2002

PATENT APPLICATION: US/10/033,325

TIME: 11:12:55

Input Set : A:\GC502-2-C1-SEQLIST.txt

Output Set: N:\CRF3\07082002\J033325.raw

```

234          165          170          175
235 Val Gly Ala Val Asp Ser Ser Asn Gln Arg Ala Ser Phe Ser Ser Val
236          180          185          190
237 Gly Pro Glu Leu Asp Val Met Ala Pro Gly Val Ser Ile Gln Ser Thr
238          195          200          205
239 Leu Pro Gly Asn Lys Tyr Gly Ala Tyr Asn Gly Thr Ser Met Ala Ser
240          210          215          220
241 Pro His Val Ala Gly Ala Ala Ala Leu Ile Leu Ser Lys His Pro Asn
242          225          230          235          240
243 Trp Thr Asn Thr Gln Val Arg Ser Ser Leu Glu Asn Thr Thr Thr Lys
244          245          250          255
245 Leu Gly Asp Ser Phe Tyr Tyr Gly Lys Gly Leu Ile Asn Val Gln Ala
246          260          265          270
247 Ala Ala Gln
248          275
250 <210> SEQ ID NO: 4
251 <211> LENGTH: 275
252 <212> TYPE: PRT
253 <213> ORGANISM: Bacillus subtilis
255 <400> SEQUENCE: 4
256 Ala Gln Ser Val Pro Tyr Gly Ile Ser Gln Ile Lys Ala Pro Ala Leu
257 1          5          10          15
258 His Ser Gln Gly Tyr Thr Gly Ser Asn Val Lys Val Ala Val Ile Asp
259          20          25          30
260 Ser Gly Ile Asp Ser Ser His Pro Asp Leu Asn Val Arg Gly Gly Ala
261          35          40          45
262 Ser Phe Val Pro Ser Glu Thr Asn Pro Tyr Gln Asp Gly Ser Ser His
263          50          55          60
264 Gly Thr His Val Ala Gly Thr Ile Ala Ala Leu Asn Asn Ser Ile Gly
265          65          70          75          80
266 Val Leu Gly Val Ser Pro Ser Ala Ser Leu Tyr Ala Val Lys Val Leu
267          85          90          95
268 Asp Ser Thr Gly Ser Gly Gln Tyr Ser Trp Ile Ile Asn Gly Ile Glu
269          100          105          110
270 Trp Ala Ile Ser Asn Asn Met Asp Val Ile Asn Met Ser Leu Gly Gly
271          115          120          125
272 Pro Thr Gly Ser Thr Ala Leu Lys Thr Val Val Asp Lys Ala Val Ser
273          130          135          140
274 Ser Gly Ile Val Val Ala Ala Ala Ala Gly Asn Glu Gly Ser Ser Gly
275          145          150          155          160
276 Ser Thr Ser Thr Val Gly Tyr Pro Ala Lys Tyr Pro Ser Thr Ile Ala
277          165          170          175
278 Val Gly Ala Val Asn Ser Ser Asn Gln Arg Ala Ser Phe Ser Ser Ala
279          180          185          190
280 Gly Ser Glu Leu Asp Val Met Ala Pro Gly Val Ser Ile Gln Ser Thr
281          195          200          205
282 Leu Pro Gly Gly Thr Tyr Gly Ala Tyr Asn Gly Thr Ser Met Ala Thr
283          210          215          220
284 Pro His Val Ala Gly Ala Ala Ala Leu Ile Leu Ser Lys His Pro Thr

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/033,325

DATE: 07/08/2002

TIME: 11:12:56

Input Set : A:\GC502-2-C1-SEQLIST.txt

Output Set: N:\CRF3\07082002\J033325.raw